Engineering Biology

Research Article

IET Journals
The Institution of Engineering and Technology

Genome-scale model of *C. autoethanogenum* reveals optimal bioprocess conditions for high-value chemical production from carbon monoxide

eISSN 2398-6182
Received on 9th May 2018
Revised 5th March 2019
Accepted on 20th March 2019
doi: 10.1049/enb.2018.5003
www.ietdl.org

Rupert O.J. Norman[1,2] ✉, Thomas Millat[1], Sarah Schatschneider[1,3], Anne M. Henstra[1], Ronja Breitkopf[1], Bart Pander[1], Florence J. Annan[1], Pawel Piatek[1], Hassan B. Hartman[4,5], Mark G. Poolman[4], David A. Fell[4], Klaus Winzer[1], Nigel P. Minton[1], Charlie Hodgman[1,2]

[1]*Synthetic Biology Research Centre, University of Nottingham, University Park, Nottingham NG7 2RD, UK*
[2]*School of Biosciences, University of Nottingham, Sutton Bonington Campus, Sutton Bonington, Leicestershire LE12 5RD, UK*
[3]*Evonik Nutrition and Care GmbH, Kantstr. 2, 33798 Halle-Kinsbeck, Germany*
[4]*Department of Biology and Medical Sciences, Oxford Brookes University, Oxford OX3 0BP, UK*
[5]*Public Health England, 61 Colindale Avenue, London NW9 5EQ, UK*
✉ *E-mail: rupert.norman@nottingham.ac.uk*

Abstract: *Clostridium autoethanogenum* is an industrial microbe used for the commercial-scale production of ethanol from carbon monoxide. While significant progress has been made in the attempted diversification of this bioprocess, further improvements are desirable, particularly in the formation of the high-value platform chemicals such as 2,3-butanediol (2,3-BD). A new, experimentally parameterised genome-scale model of *C. autoethanogenum* predicts dramatically increased 2,3-BD production under non-carbon-limited conditions when thermodynamic constraints on hydrogen production are considered.

1 Introduction

The adverse environmental and societal consequences of continued fossil-fuel dependence represent arguably *the* defining challenge for scientific research in the 21st century [1, 2]. One component of the solution to this problem is carbon recycling, which can be achieved through the application of bioprocesses enabling conversion of industrial waste-gas components into commodity chemicals [3]. For example, the Chicago-based company LanzaTech (www.lanzatech.com) uses *Clostridium autoethanogenum* to convert carbon monoxide (CO) into ethanol on a commercial scale [4–6]. Diversification of the product portfolio of this sustainable technology will help to secure its continued success.

The fundamental metabolic pathway enabling *C. autoethanogenum* to grow on CO is the Wood–Ljungdahl pathway (WLP) [7, 8]. This ancient biochemical pathway is split into two 'branches': the methyl branch and the carbonyl branch [9]. The methyl branch proceeds by constructing a methyl group from carbon dioxide (CO_2) *via* a series of biochemical conversions including an adenosine triphosphate (ATP)-consuming reaction [catalysed by formyl-tetrahydrofolate (THF) ligase, FtfL] and three redox reactions collectively requiring nicotinamide adenine dinucleotide (NADH), NAD phosphate (NADPH) and reduced ferredoxin (Fd_{red}). The carbonyl branch is simply the reduction of CO_2 to an enzyme-bound carbonyl group as catalysed by the acetyl-CoA synthase/CO dehydrogenase complex (ACS/CODH). This same enzyme complex is responsible for the final step of the pathway, in which the methyl group and enzyme-bound carbonyl group are combined with CoA to form one molecule of acetyl-CoA.

Clearly, in the case of CO-fed growth, sources of CO_2, reducing power and the means to generate ATP are needed for the WLP to function. As with all known acetogens [10], acetate and ATP are produced from acetyl-CoA *via* phosphotransacetylase and acetate kinase, thus fulfilling the energetic requirement of FtfL but not enabling a positive ATP yield [5]. The CO_2 and reducing power requirements are met by two additional monofunctional CODH enzymes encoded in the *C. autoethanogenum* genome which generate Fd_{red} and CO_2 from CO (and water). Fd_{red} is then oxidised to generate NADH and subsequently NADPH, which in turn satisfy the redox requirements of the WLP. Finally, electrons are transferred from Fd_{red} to NAD by a membrane-bound Rnf complex [11], which produces a transmembrane proton gradient and enable the generation of ATP by F_1F_o ATP synthase [12, 13]. The regeneration of NAD for this essential process is achieved through further redox reactions leading to excreted by-products, *e.g.* ethanol [5].

One such product of particular importance is the platform chemical 2,3-butanediol (2,3-BD) – a trace component of the native *C. autoethanogenum* product profile, the downstream products of which have an estimated global market value of $43 billion in sales [14]. Other compounds in the native product profile of *C. autoethanogenum* include acetate, ethanol and, in smaller amounts, lactate [14, 15]. Hydrogen gas is another common product in carboxydotrophic microorganisms [16]. While the potential for 2,3-BD synthesis from industrial waste gas is well-established [14], a clear bioprocess/metabolic engineering strategy for its increased production has not been published despite the construction and analysis of a genome-scale model (GSM) of *C. autoethanogenum* [17] – an approach which has achieved success in guiding metabolic engineering through the investigation of optimal steady states [18]. Previous attempts to model *C. autoethanogenum* use parameters from GSMs of other organisms [17, 19, 20], a potential weakness identified by Dash *et al.* [21]. The availability of a new, manually annotated genome sequence of *C. autoethanogenum* [22], therefore, provides an opportunity to construct an improved, experimentally parameterised GSM and pursue novel strategies for enhanced 2,3-BD production.

Experimental studies aimed at manipulating the product profile of *C. autoethanogenum* have so far focused on carbon-limited growth regimes [17, 19], *i.e.* the scenario, in which rate of growth of a bacterial population is limited by the supply of its carbon source. Such studies tend to neglect non-carbon limitation – the case where some other nutrient becomes limiting and the carbon source is in excess. We argue that non-carbon-limited growth

Table 1 Genome sequences for *Clostridia* spp. referred to throughout this work

Organism	Year	Reference	Annotation	Structure
C. autoethanogenum	2013	[27]	automated	100 contigs
	2014	[24]	automated	whole genome
	2015	[22]	manual	whole genome
C. ljungdahlii	2010	[28]	automated	whole genome

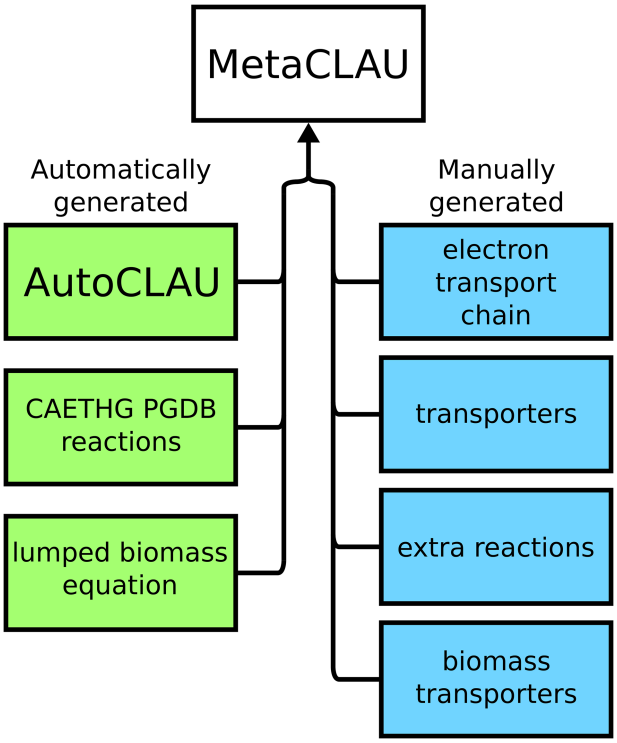

Fig. 1 *ScrumPy model structure for MetaCLAU. Coloured boxes represent the different 'modules' referenced by the top-level file, 'MetaCLAU.spy': Green – automatically generated modules, blue – manually defined modules. The arrow illustrates a compilation of modules to form MetaCLAU*

regimes are worth investigating because an excess supply of energy (*i.e.* electrons from CO) may result in the diversion of flux toward increasingly reduced compounds (*e.g.* 2,3-BD), especially in strict anaerobes such as *C. autoethanogenum* whose product profile is constrained by the redox balance of internal metabolites [23].

This work presents the construction of a new GSM of *C. autoethanogenum* ('MetaCLAU') including parameterisation and validation with data from continuous-culture chemostat experiments. The subsequent model is used to test the hypotheses that 2,3-BD is an optimal sink for excess reducing power supplied under non-carbon-limited conditions before and after a thermodynamic limit on hydrogen production is imposed.

2 Methods

2.1 Data

This section describes the data used for model construction, parameterisation and validation. Data used in the interpretation of model analyses are also specified.

2.1.1 Genome sequences: The task of constructing a GSM requires an annotated genome sequence. Table 1 summarises relevant genomic data available for use in this work. The ScrumPy-readable pathway genome database (PGDB) generated from the genome sequence described in [24] is available from www.metacyc.org. Flatfiles for the PGDB derived from the genome annotation described in [22] were generated with Pathway Tools [25] and can be found in [26].

2.1.2 Fermentation data: Optical density (OD), gas chromatography and cell-mass data collected at steady state during fermentation experiments were used to derive model parameters in this work. Bioreactor data were collected and processed using in-house software (see 2.7). Details of fermentation experiments are provided in the supplementary materials (S4).

The data for the estimation of growth and non-growth-associated maintenance costs are shown in the supplementary materials (S6).

2.1.3 Biomass: Proportions of macromolecular biomass components were measured experimentally with continuous-culture samples taken at steady state. Details of experimental procedures for the measurement of total DNA, RNA, protein, lipid and polysaccharide are given in the supplementary materials. Liquid chromatography-mass spectrometry (LC-MS) was used to estimate trace metabolite concentrations as described in [29].

2.2 Model construction

In the context of this paper, 'construction' refers to the selection of a set of reaction stoichiometries which form the metabolic network, whereas 'curation' refers to the identification and correction of errors and omissions in the definition of these reactions, *e.g.* errors in mass-balance and thermodynamic favourability.

ScrumPy models are formed of a top-level 'module' and several sub-modules containing either automatically generated or manually defined reaction stoichiometries (see Fig. 1). This helps to organise the model components during construction and curation.

2.2.1 Draft network: Construction of the GSM began with the Tier 3 BioCyc [25, 30] database for *C. autoethanogenum* JA1-1 (strain DSM 10061) generated from the genome sequence published by Brown *et al.* [24]. This database (referred to here as the 'CAETHG database') formed the foundation of a draft reconstruction of the organism's metabolic network. Automatically generated Tier 3 BioCyc databases are not fully curated, thus an additional manual genome annotation [22] was required to complete construction in line with methods described by Fell *et al.* [31] and Hartman *et al.* [32]. This additional annotation was used to create a second genome database (referred to as the 'CLAU database') with the PathoLogic algorithm as implemented in Pathway Tools [25, 30]. Information contained in this second database formed the basis for the continued curation of the model, the aim being to establish gene-protein-reaction relationships (GPRs) based solely on the manual annotation. For a detailed, step-by-step model construction methodology see supplementary materials (S1).

2.2.2 Curation: Initial curation steps involved removing chemically undefined metabolites from the automatically generated network. Subsequently, reactions without gene associations were investigated and removed if evidence for the necessary encoded enzymes could not be found [22]. Atomically unbalanced reactions were also corrected or removed. The thermodynamic consistency of the model, *i.e.* its adherence to energy conservation, was regularly checked using a specific linear programming (LP) problem detailed in Section 2.2.4. Inconsistent enzyme subsets containing erroneous reaction reversibility constraints [33] were identified using ScrumPy [34] and removed from the model (see Section 2.5).

2.2.3 Electron transport chain: A model of the electron transport chain (ETC) of *C. autoethanogenum* was manually constructed based on biochemical literatures [9, 11, 14, 35, 36]. The separate construction of an ETC sub-network enables the computation of its elementary modes (*i.e.* the set of minimal steady-state flux distributions across the network [37, 38]), since smaller networks avoid combinatorial difficulties encountered with larger networks. Inspection of the subsequent elementary modes ensures that legitimate routes for the generation of energy are available in model simulations (see Fig. 2).

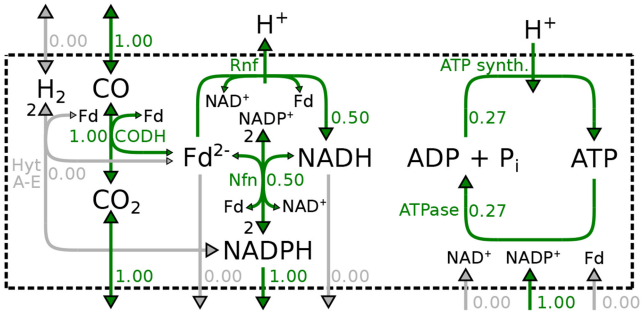

Fig. 2 *Elementary mode of the ETC. Green arrows represent active reactions, grey arrows represent inactive reactions. Reversible reactions are represented as bi-directional arrows. Relative flux values and enzyme names are written next to arrow lines. Flux carried by irreversible reactions proceed in the direction indicated by the arrowhead, whereas reversible reactions carry flux in the 'downward' direction (in terms of the figure's orientation)*

One feature of the ETC in *C. autoethanogenum* is that the number of translocated protons required for the generation of one ATP by ATP synthase is not known. In previous studies, a value of 3.66 H^+/ATP is assumed based on the structure of the *Clostridium paradoxum* ATP synthase (the rotor of which consists of 11 c-subunits) [39], taking this stoichiometry as an indication of the *Clostridium* phenotype, in general [35]. We adopt this assumption here but also show that our modelling results are not qualitatively sensitive to this parameter [see supplementary materials (S3)].

The details of the ETC are as follows: *C. autoethanogenum* maintains a transmembrane proton gradient enabling energy conservation *via* coupling of an Rnf complex (1) with an ATP synthase (2) [11]. When an ATPase c-ring consisting of 11 c-subunits is assumed (as in [35]), the following reactions represent the ETC:

$$\mathrm{Rnf}: 3H^+ + NAD^+ + Fd_{red} \rightarrow 2H^+_{ext} + NADH + Fd_{ox} \quad (1)$$

$$\mathrm{ATP\,synth}: 3.67H^+_{ext} + ADP + P_i \rightarrow 2.67H^+ + ATP + H_2O \quad (2)$$

$$\mathrm{CODH}: CO + Fd_{ox} + H_2O \rightarrow 2H^+ + CO_2 + Fd_{red} \quad (3)$$

$$\mathrm{MetFV}: \mathrm{methylene-THF} + 2NADH + Fd_{ox} \rightarrow \mathrm{methyl-THF} + 2NAD^+ + Fd_{red} \quad (4)$$

The stoichiometry of the reaction catalysed by methylene-THF reductase (MetFV) is unknown [40]. The definition given above (4) represents the electron-bifurcating mechanism proposed by Köpke *et al.* [28] and further developed in [35]. An electron-bifurcating hydrogenase, HytA-E is responsible for the utilisation of hydrogen as an energy source [36] and is also thought to be responsible for hydrogen production [41] (5). Finally, since the WLP in *C. autoethanogenum* involves an NADPH-dependent methylene-THF dehydrogenase, the ETC must include a mechanism for the generation of NADPH (this is also necessary to support biomass). This is achieved by the Nfn complex (6)

$$\mathrm{HytA-E}: 2H_2 + NADP^+ + Fd_{ox} \rightarrow 3H^+ + NADPH + Fd_{red} \quad (5)$$

(see (6))

An example steady-state flux distribution of the ETC which yields ATP and NADPH is shown in Fig. 2.

2.2.4 Energy conservation: The following analysis ensured thermodynamic consistency across the metabolic network: flux balance analysis (FBA) was computed with fixed positive ATP demand and without uptake of any carbon, nutrient or energy source in line with methods described in [32]:

$$\text{minimise: } |\boldsymbol{v}| \quad \text{subject to} \begin{cases} \boldsymbol{Nv} = \boldsymbol{0} \\ v_{\mathrm{ATPase}} = 1 \\ v_{t_i} \le 0; \quad i \in \{1, \ldots, X_{\mathrm{trans}}\} \end{cases} \quad (7)$$

where transport reactions are denoted v_{t_i} and $\boldsymbol{t}$ is a vector of length X_{trans} containing all indices of transport reactions in $\boldsymbol{v}$. Any subsequent solution included a thermodynamic error, meaning definitions for one or more of the participant reactions required amendment or removal.

2.2.5 Transporters: The module 'Transporters.spy' defines all reactions involving transfer between the organism and the environment. Transporters were added based on known carbon/energy sources and products of *C. autoethanogenum* [15]. Transporters were also added for individual biomass precursors [31] and components of the growth media used to culture *C. autoethanogenum* in this paper, see supplementary materials (S4).

2.3 Model characteristics

A GSM of *C. autoethanogenum* consisting of 755 reactions and 772 metabolites has been constructed through the refinement of a draft network derived from a Pathway Tools database consisting of 1429 reactions and 1097 metabolites. About 73 model reactions in the curated model were unique to the initial CAETHG database. About 47 gene associations for these reactions were mapped to CLAU locus tags using EDGAR [42]. About 15 of the remaining model reactions unique to the CAETHG database were defined as spontaneous, and thus required no association to genes. Gene associations for 3 of the remaining 11 reactions were defined manually, and 8 were retained as they proved essential for the production of biomass on a range of growth substrates (CO, [CO + H_2] and fructose). These reactions, forming a pathway for teichoic acid production, have been retained as hypothetical reactions and are detailed in the supplementary material (S1). Teichoic acid is an assumed component of the model's biomass equation consistent with established physiology of Gram-positive bacteria [43].

2.4 Parameterisation

Where possible, species-specific parameters should be derived experimentally, improving the accuracy of model calculations and discouraging the unhelpful culture of 'parameter borrowing'. In the case of bacterial cells, two key parameters are required: biomass composition and ATP maintenance costs.

2.4.1 Biomass composition: A GSM should be capable of producing essential cellular materials (in realistic proportions) using feasible biosynthetic routes to enable increasingly accurate calculations of optimal network behaviours [44, 45]. Despite a large number of biotechnologically interesting *Clostridia*, detailed physiological information suitable for deriving cellular biomass composition is available for a small range of species only [46–48].

The macromolecular biomass composition was measured experimentally including protein, DNA, RNA, lipid and polysaccharide [see supplementary materials (S4) for experimental methods]. The relative abundances of various lipids in *C. autoethanogenum* are taken from [15]. DNA nucleotide ratios were estimated from the full genome sequence [22], whereas nucleotide and amino acid ratios were estimated from transcribed rRNA and translated ribonucleoprotein sequences, respectively [see supplementary materials (S4)]. The use of ribosomal RNA and protein sequences to characterise the prevalence of ribonucleotides and amino acids are based on the observation that the largest allocation of protein synthesis capacity is in the production of ribosomes [49]. The 'Other' class of biomass component contains

$$\mathrm{Nfn}: 2NADP^+ + NADH + H^+ + Fd_{red} \rightarrow 2NADPH + NAD^+ + Fd_{ox} \quad (6)$$

Table 2 Overall biomass composition of *C. autoethanogenum* DSM 10061

Component	Percentage	±
protein	26.3	2.3
DNA	14.6	7.5
RNA	17.9	4.2
lipid	21.0	1.7
polysaccharide	10.2	7.8
teichoic acid	0.6	0.1
other	9.4	—

Table 3 Maintenance costs (r_{ATP}) calculated at $\mu = 0.028\,h^{-1}$ with GAM and NGAM values from a range of GSMs/organisms

Organism	GAM	NGAM	r_{ATP}	Reference
C. autoethanogneum	111.2	2.2	5.3	—
C. autoethanogneum	41.3	—	1.2	[17]
C. ljungdahlii	46.7	0.45	1.8	[50]
Clostridium beijerinckii	40.0	8.5	9.6	[51]
Acetobacter woodii	150.0	0.8	5.0	[52]
Bacillus subtilis	105.0	9.0	11.9	[53]
Escherichia coli	97.0	18.9	21.6	[54]
Saccharomyces cerevisiae	71–91	<1	<3.3	[55]

Units for growth-associated and non-growth-associated maintenance are mmol $gDCW^{-1}$ and mmol $gDCW^{-1}\,h^{-1}$, respectively.

THF, CoA, flavin adenine dinucleotide, flavin mononucleotide, NAD and NADP, see supplementary materials (S4).

The overall biomass composition including essential trace metabolite pools is given in Table 2. Tables showing the estimated relative proportion of each biomass precursor within the macromolecular divisions are given in the supplementary materials (S4).

2.4.2 ATP maintenance costs: The following experimental procedure was conducted to measure growth and non-growth ATP maintenance costs (growth-associated ATP maintenance cost (GAM) and non-growth-associated ATP maintenance cost (NGAM), respectively) in CO-fed continuous culture of *C. autoethanogenum*: The set dilution rate D was increased incrementally from 0.01 to $0.028\,h^{-1}$. Steady state was reached after each change, as determined by stable culture OD. For each steady-state period, the average CO uptake rate was calculated and plotted against the corresponding specific growth rate μ which was taken as equivalent to the set dilution rate according to (8). The slope of a linear fit to this data is the growth-associated CO uptake rate; the y-intercept is the non-growth-associated CO uptake rate [see supplementary materials (S6)]

$$\frac{\mathrm{d(OD)}}{\mathrm{d}t} = (\mathrm{OD})(\mu - D) = 0 \tag{8}$$

GAM and NGAM were calculated by multiplying growth and non-growth-associated CO uptake rates by the model-calculated maximal ATP yield (Y_{ATP}) from CO. The growth and non-growth-associated CO uptake rates are 415 ± 92 and 5.7 ± 1.4, respectively, and the optimal Y_{ATP} value is 0.375 [see (12)]. These values give a GAM of 155.7 ± 34.6 and an NGAM value of 2.2 ± 0.5 (The units for GAM and NGAM are mmol $gDCW^{-1}$ and mmol $gDCW^{-1}\,h^{-1}$, respectively.).

Since a proportion of GAM is accounted for by biomass precursor production in the metabolic network, this proportion must be subtracted from the total GAM before incorporation as a flux constraint. After subtracting the proportion of growth-associated ATP maintenance incorporated in the model's metabolic network (44.5 mmol $gDCW^{-1}$) from the experimentally derived GAM parameter, the actual GAM value used as a constraint in FBA simulations was 111.2 mmol $gDCW^{-1}\,h^{-1}$. Total ATP maintenance (r_{ATP}) resulting from GAM/NGAM parameters sourced from published GSMs and an assumed growth rate of $0.028\,h^{-1}$ are shown in Table 3.

2.5 Steady-state metabolism

The fundamental concept underlying the analyses presented in this work is metabolic steady state. This occurs when the concentrations of metabolites involved in a metabolic system do not change over time [56, 57]. The steady-state constraint is enforced with a set of linear differential equations representing the production and consumption of m metabolites by n chemical reactions. These equations can be represented as the dot product of a matrix of reaction stoichiometries $\boldsymbol{N}$ and a vector of reaction fluxes (*i.e.* net reaction rates) $\boldsymbol{v}$ [58]. The steady-state condition is fulfilled when

$$\frac{\mathrm{d}\boldsymbol{m}}{\mathrm{d}t} = \boldsymbol{N}\boldsymbol{v} = 0, \tag{9}$$

i.e. when the distribution of metabolic fluxes $\boldsymbol{v}$ lies in the nullspace of $\boldsymbol{N}$. Basis vectors for the nullspace of $\boldsymbol{N}$ can subsequently be analysed to reveal useful properties of the subject metabolic network such as 'dead reactions' which always carry zero flux at steady state [59]. Another useful concept from the nullspace of $\boldsymbol{N}$ is 'enzyme subsets' [60]. An enzyme subset is a subset of enzymes within a metabolic network whose members carry flux in a fixed ratio at steady state [59]. Inconsistent enzyme subsets contain one-or-more enzymes whose flux direction violates irreversibility constraints, effectively inactivating every enzyme in the subset.

2.6 Flux balance analysis

FBA is a fundamental technique for the analysis of GSMs which formulates steady-state metabolism as an LP problem given some assumed, biologically relevant objective function [61–63].

Since (9) only specifies constraints on the relative values of fluxes, further constraints are required to achieve a solution space which is sufficiently bounded for LP [64]. For example, transport 'reactions' responsible for the influx (and efflux) of metabolites into the system are often subject to an upper-bound constraint. Following this, an objective function is specified which maximises or minimises a selection of the flux variables. An objective is usually chosen which mimics an assumed biological objective [65] such as maximising ATP generation. The subsequent solution to the LP problem is thus an optimal flux distribution.

Results obtained through the application of FBA provide insight into the capabilities of a metabolic network, thus helping to investigate the steady-state behaviour of an organism. In this paper, FBA is used to calculate the growth rate of *C. autoethanogenum* as a means of model validation. Following this, FBA is used to investigate the model's response to changes in industrially relevant culture conditions including non-carbon limitation and constraints on hydrogen production.

2.6.1 Network flux minimisation: The preferred objective function in this paper is the minimisation of absolute flux through all enzyme-catalysed reactions (10). For a mass-balanced metabolic network, total network flux minimisation reduces the occurrence of thermodynamically infeasible energy-generating cycles, while providing a reasonable approximation of enzymatic economy [31, 32, 34, 66, 67].

$$\begin{aligned} &\text{minimise: } |\boldsymbol{v}_{cat}| \quad \text{(objective function)} \\ &\text{subject to} \begin{cases} \boldsymbol{N}\boldsymbol{v} = \boldsymbol{0} \quad \text{(steady state)} \\ v_{ATPase} = Y_{xATP} \times \mu + m_{ATP} \\ v_{k_i} = \mu \times b_i; \quad i \in \{1, \ldots, X_{bio}\} \\ v_{p_j} \le 0; \quad j \in \{1, \ldots, X_{prods}\} \end{cases} \end{aligned} \tag{10}$$

Table 4 Comparison of genome-scale metabolic models

Model	Reactions	Dead reactions	Trans.	Mets	Genes	GPRs	IESs*	Energy conserved?
MetaCLAU	755	327	80	772	699	600	0	YES
iCLAU786 [17]	1001	810	28	1046	802	841	21	YES
iCLAU786 [20]	1109	966[a]	66	1162	784	862	22	**NO**
CarveMe	1202	30	129	922	809	844	15	YES
iHN637	785	186	95	698	637	615	20	YES

[a]This high proportion of dead reactions is due to the inclusion of external metabolites in the stoichiometric matrix of iCLAU786. See externally hosted files [26].

*Inconsistent enzyme subsets (IESs)

where $\boldsymbol{N}$ is the stoichiometric matrix and $\boldsymbol{v}$ is a vector of all reaction flux values included in the metabolic network; $\boldsymbol{v}_{\text{cat}}$ is the set of network reactions catalysed by enzymes, *i.e.* the subset of $\boldsymbol{v}$ which excludes transport by diffusion (in this paper: CO, CO_2 and H_2) and spontaneous reactions. This objective requires information on whether or not a reaction can occur spontaneously in the context of metabolism, as provided in Pathway Tools. $Y_{x\text{ATP}}$ and m_{ATP} are equivalent to the growth and non-growth-associated ATP maintenance costs (GAM and NGAM), respectively, and μ is the specific growth rate. v_{ATPase} is the flux through the network's ATPase reaction ($\text{ATP} + \text{H}_2\text{O} \rightarrow \text{ADP} + P_i + \text{H}^+$). Reaction fluxes representing the production of X_{bio} biomass precursors are denoted v_{k_i}, where $\boldsymbol{k}$ is a vector containing the corresponding indices of these reactions in $\boldsymbol{v}$. The vector $\boldsymbol{b}$ contains relative abundance values for each biomass precursor which, when multiplied by μ, determine the value of each v_{k_i}. Reaction fluxes representing the transport of non-biomass-associated metabolites not required as growth substrates (*i.e.* potential products) are denoted v_{p_j}, where $\boldsymbol{p}$ is a vector of length X_{prods} containing the corresponding indices in $\boldsymbol{v}$.

2.6.2 Growth yield/ATP maximisation: An alternative objective function is the maximisation of biomass precursor production, an approximation of growth rate optimisation [44]

$$\text{maximise}: \mu$$
$$\text{subject to} \begin{cases} \boldsymbol{N}\boldsymbol{v} = \boldsymbol{0} \\ v_{\text{ATPase}} = m_{\text{ATP}} \\ v_{k_i} = g_i; \quad i \in \{1, \ldots, X_{\text{growth}}\} \\ v_{p_j} \leq 0; \quad j \in \{1, \ldots, X_{\text{prods}}\} \end{cases} \tag{11}$$

Here, μ is represented as a lumped chemical equation including all biomass precursors as reactants with their relative abundance given as stoichiometric coefficients. $Y_{x\text{ATP}}$ (GAM) is incorporated in the lumped equation as the stoichiometric coefficient on ATP. Reaction fluxes representing each of X_{growth} substrate-uptake reactions are denoted v_{k_i}, where $\boldsymbol{k}$ is a vector containing the corresponding indices in $\boldsymbol{v}$. The vector $\boldsymbol{g}$ contains fixed-flux values for each of X_{growth} substrate-uptake reactions. In this paper, the fixed substrate-uptake rate is taken as the average experimentally observed uptake rate over the time period at which the culture is at steady state.

An alternative form of this LP problem maximises ATP dissipation, *i.e.* flux carried by ATPase (v_{ATPase}) in the ATP hydrolysing direction. Solutions to this problem achieve optimal ATP yields. Furthermore, a complete hierarchy of products ordered in terms of Y_{ATP} can be generated by iteratively computing the Y_{ATP}-optimal solution while blocking (*i.e.* constraining to 0) flux across the transporter of the product formed in the previous iteration until ATP production is infeasible.

2.7 Software

Reconstruction steps and model analyses were carried out using the *ScrumPy* package (http://mudshark.brookes.ac.uk/ScrumPy) [34]. Analysis of the previously published iCLAU786 [17] model was carried out in COBRApy [68]. Bioreactor data were analysed using an automated python/MATLAB-based software pipeline designed for the BioCommand software (BioCommand®, New Brunswick Scientific), in which data were transferred to a web server, allowing real-time updates before calibration and post-processing using a suite of MATLAB® scripts. Further details of this toolset including source code and descriptions are available on request.

3 Results

3.1 Comparison of MetaCLAU with published models

Genome-scale metabolic models have been published for both *Clostridium ljungdahlii* [50] and *C. autoethanogenum* [17, 20]. The *C. autoethanogenum* model, iCLAU786, was first made available in [17] and then updated in [20]. Additionally, the repository of metabolic models automatically generated by CarveMe [69] contains a draft network for *C. autoethanogenum*. This section shows the results of a general comparison between MetaCLAU and these published models.

Differences in the general model statistics are shown in Table 4. Importantly, the more recent iCLAU786 version [20] fails the essential condition set-out in Section 2.2.4 that no ATP can be generated without the uptake and excretion of metabolites. For this reason, this version of iCLAU786 is excluded from further analysis. Similarly, the CarveMe draft model is excluded from further analysis as it is unable to generate ATP from CO and water.

An important difference is the number of dead reactions; iCLAU786 has 966 (97%) dead reactions, whereas MetaCLAU has 327 (43%). MetaCLAU also accounts for the transport of more metabolites. Finally, unlike the other models, MetaCLAU contains no inconsistent enzyme subsets [33]. Details of this comparison are provided in the externally hosted files [26].

3.2 Validation

Validation of key behaviours against experimental data was carried out to demonstrate the model's good reliability. Substrate utilisation, the growth rate on CO and product spectrum are the validating behaviours tested in this paper.

3.2.1 Substrate utilisation: The ability of *C. autoethanogenum* to use different carbon/energy sources for the production of biomass (and maintenance) was calculated with FBA. Each compound was used as the sole carbon source in the *in silico* minimal medium by allowing influx across the corresponding transporter while applying the min $|\boldsymbol{v}|$ objective. The simulation results were compared with experimental data (Table 5), showing good agreement with known carbon/energy sources.

3.2.2 Growth rate: An average experimentally observed CO uptake rate of 16.52 ± 0.02 mmol $\text{gDCW}^{-1}\,\text{h}^{-1}$ was applied as a fixed-flux constraint on CO transport, which, with the objective function set to maximise biomass production, resulted in a calculated optimal specific growth yield of $0.027\,\text{h}^{-1}$. This value is within the error margin of the measured specific growth rate of $0.028 \pm 0.001\,\text{h}^{-1}$. A comparison of model-simulated *versus* reported experimental growth rates including fructose and syngas is shown in Table 6.

3.2.3 Product spectrum: The ratio of product fluxes in optimal solutions with CO as the sole source of carbon and energy was sensitive to the choice of the objective function. Maximisation of growth yield (11) resulted in acetate and CO_2 production, with no flux to ethanol. In contrast, when minimisation of network flux was the objective function (10), acetate and ethanol formed the main products. Optimal flux distributions computed with published models of *C. autoethanogenum* (iCLAU786 [17, 19]) and *C. ljungdahlii* (iHN637 [50]) did not include simultaneous production of ethanol and acetate without additional constraints (see Table 7). The min $|\boldsymbol{v}_{cat}|$ objective could not be applied to iCLAU786 or iHN637 since reactions that are known to occur spontaneously (in the context of metabolism) are not identified in these models.

The optimal ATP yield (Y_{ATP}) per mole substrate uptake with CO as the sole carbon and energy source was 0.375 with acetate as the sole product [net stoichiometry shown in (12)]. A sub-optimal Y_{ATP} value (0.342) is associated with ethanol production. A hierarchy of products based on their associated ATP yield is shown below, demonstrating that the model can return flux distributions associated with the expected range of liquid products:

$$Y_{ATP}^{acetate} = 0.375\,\mathrm{CO} + 1/2\mathrm{H_2O} \rightarrow 1/4\mathrm{CH_3COO^-} + 1/4\mathrm{H^+} + 1/2\mathrm{CO_2} \quad (12)$$

$$Y_{ATP}^{ethanol} = 0.341\,\mathrm{CO} + 1/2\mathrm{H_2O} \rightarrow 1/6\mathrm{CH_3CH_2OH} + 2/3\mathrm{CO_2} \quad (13)$$

$$Y_{ATP}^{lactate} = 0.174\,\mathrm{CO} + 1/2\mathrm{H_2O} \rightarrow 1/6\mathrm{CH_3CH(OH)CO_2^-} + 1/6\mathrm{H^+} + 1/2\mathrm{CO_2} \quad (14)$$

$$Y_{ATP}^{2,3BD} = 0.14\,\mathrm{CO} + 5/11\mathrm{H_2O} \rightarrow 1/11\mathrm{CH_3(CHOH)_2CH_3} + 7/11\mathrm{CO_2} \quad (15)$$

$$Y_{ATP}^{H_2} = 0.136\,\mathrm{CO} + \mathrm{H_2O} \rightarrow \mathrm{H_2} + \mathrm{CO_2} \quad (16)$$

Table 5 Carbon-source validation results

Carbon source	Model predicts growth?	Growth observed?	Reference
CO	YES	YES	[15, 17, 35]
CO_2 (+ H_2)	YES	YES	[15, 35]
fructose	YES	YES	[15, 35]
xylose	YES	YES	[15, 70]
arabinose	YES	YES	[15]
rhamnose	YES	YES	[15]
pyruvate	YES	YES	[15]
glutamate	YES	YES	[15]
formate	YES	NO	[15]
fumarate	YES	YES[a]	[15, 71]
glucose	NO[b]	NO	[15]
lactate	NO	NO	[15]
acetate	NO	—	—
ethanol	NO	—	—
2,3-BD	YES	—	—

[a]Abrini *et al.* [15] and Breitkopf [71] disagree, see discussion.

[b]*C. autoethanogenum* does not encode a glucose transporter.

Table 6 Growth rate comparison: model simulation *versus* experimental data

Carbon/ energy source	Uptake (measured)	Growth rate (measured)	Growth rate (simulated)	Data source
CO	16.5	0.028	0.027	(this study)
CO, H_2	22.4, 13.9	0.041	0.063	[20]
fructose	1.4	0.05	0.027	[72]

3.3 Gas shift

A range of CO uptake rates exceeding the uptake rate required to support a fixed growth rate ($v_{CO} = 16.97$, $\mu = 0.028$) was applied to simulate non-carbon-limited growth. FBA with CO uptake in the range 17–42.5 mmol $gDCW^{-1}\,h^{-1}$ and a fixed growth rate of $0.028\,h^{-1}$ results in a shift from majority acetate production to ethanol production (Fig. 3) and subsequent hydrogen production in optimal solutions (Fig. 4). When the model was constrained to allow no production (or uptake) of hydrogen, 2,3-BD and lactate efflux appear in optimal solutions after the initial switch from acetate to ethanol (where 2,3-BD efflux increases with CO uptake but lactate efflux remains constant, see Fig. 5). The constraint on hydrogen efflux was applied to mimic possible limitations on hydrogen production caused by increasing hydrogen partial pressure in the external environment [73].

4 Discussion

A GSM of *C. autoethanogenum* has been constructed and parameterised with experimentally derived values representing ATP maintenance costs and cellular biomass composition. The experimental parameterisation was designed to make the model more accurately represent *C. autoethanogenum* metabolism, as confirmed by the validation in Section 3.2 above. Model-calculated yields for growth on CO agree with experimentally observed growth rate data, while the production of acetate and ethanol in optimal flux distributions supporting biomass production and maintenance costs is consistent with the established native product profile of *C. autoethanogenum*.

Oversupply of CO to the metabolic network leads to excess reducing equivalents, forcing flux through pathways (see Fig. 6) allowing the removal of superfluous reducing power (electrons) from the system [14]. This has been tested using FBA by increasing the ratio between energy-substrate supply (in this case, CO) and growth rate over a fixed range (Fig. 3) representing non-carbon-limited conditions. Results show that increasing the influx

Table 7 Ethanol, acetate and CO_2 effluxes in optimal flux distributions for CO growth computed with *C. autoethanogenum* GSMs

Model	Objective	$v_{acetate}$	$v_{ethanol}$	v_{CO_2}	CO uptake
MetaCLAU	min $\|\boldsymbol{v}_{cat}\|$	**0.06**	**2.7**	12	18.4
	min $\|\boldsymbol{v}\|$	**0.07**	**2.7**	12	18.4
	max μ	2.3	0	5	10
iCLAU786 [17]	min $\|\boldsymbol{v}\|$	0	2.4	10.9	16.7
	max μ	2.3	0	5.0	10
iHN637 [50]	min $\|\boldsymbol{v}\|$	3.1	0	8.0	15.6
	max μ	3.5	0	10.4	20

Acetate and ethanol efflux values are emboldened in flux distributions where both are produced. iCLAU786 max μ values taken from [17].

Bold values emphasise the inclusion of both acetate and ethanol in flux distributions computed with MetaCLAU.

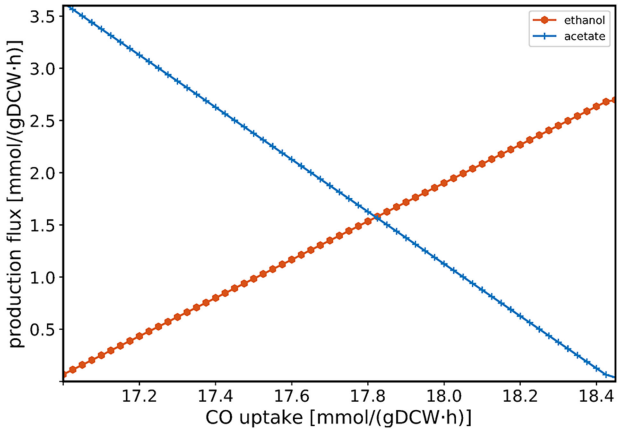

Fig. 3 *FBA-simulated gas shift in C. autoethanogenum: acetate/ethanol switch. Orange and blue lines, respectively, denote ethanol and acetate efflux*

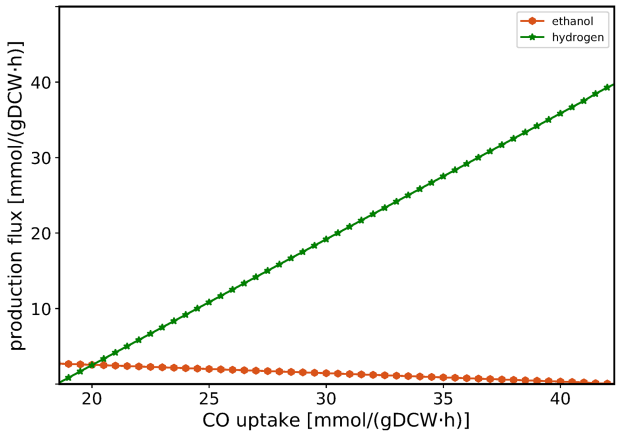

Fig. 4 *FBA-simulated gas shift in C. autoethanogenum: hydrogen production. Orange and green lines, respectively, denote ethanol and hydrogen efflux*

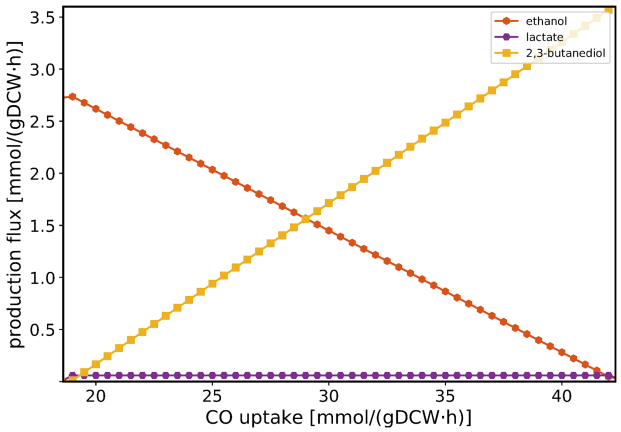

Fig. 5 *FBA-simulated gas shift in C. autoethanogenum: 2,3-BD and lactate production (hydrogen efflux blocked). Orange, yellow and purple lines, respectively, denote ethanol, 2,3-BD and lactate efflux*

of CO beyond the rate necessary to achieve a fixed growth rate causes a switch from acetate to ethanol production over the uptake range 17–18.5 mmol $gDCW^{-1}$ h^{-1}. This provides a new strategy for increasing ethanol production.

It has also been hypothesised that bacterial hydrogen production will be restricted by external constraints during continuous culture. More specifically, hydrogen production may become thermodynamically unfavourable if its partial pressure in the external environment exceeds some critical value. This will cause metabolism to favour the production of other electron sinks (such as lactate) [73–75]. To test this, the CO uptake scan was repeated

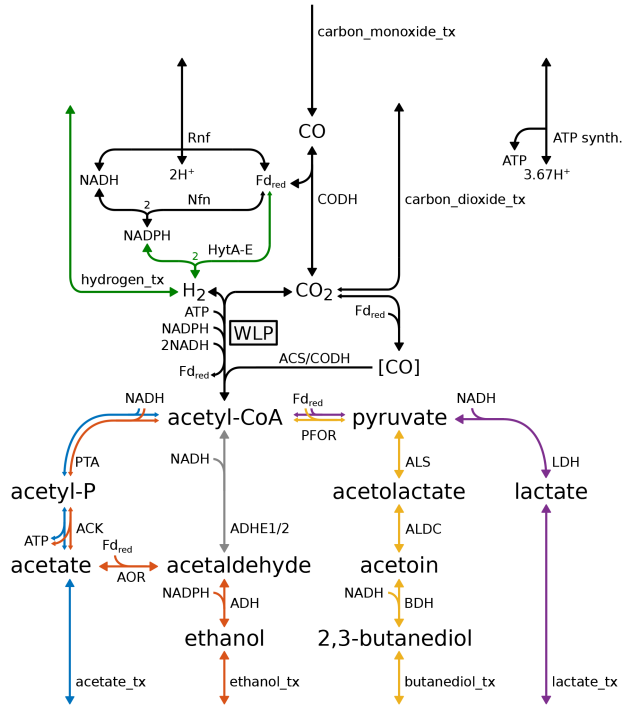

Fig. 6 *Metabolic pathways in C. autoethanogenum for CO-fed growth. Colour scheme match Figs. 3–5, showing pathways leading to corresponding products. ADHE1/2 is coloured grey to signify that it does not carry flux. Black arrows carry flux in all solutions computed during CO scans. (Only primary metabolites are shown including redox requirements of routes to acetate, ethanol, lactate and 2,3-BD. '_tx' denotes a transport process. WLP = Wood–Ljungdahl pathway. Stoichiometric weightings for Nfn and HytA-E shown)*

with hydrogen transport restricted to zero. Optimal flux distributions resulting from this simulation showed 2,3-BD production increasing monotonically with CO uptake after the acetate/ethanol switch (Fig. 5). This result is significant since 2,3-BD production has not been reported from the analysis of previous models [17, 20]. Furthermore, lactate production appeared in optimal flux distributions but was unchanged with increasing CO uptake, suggesting an association with biomass production. Thus, the analysis has uncovered a testable set of bioprocess conditions for which 2,3-BD is included in optimal metabolic flux distributions, increasing with the gas inflow.

4.1 Model availability

The model is available to download as an SBML, JSON (COBRA readable) or ScrumPy-formatted file from the Biomodels database [76] using accession number MODEL1810120001. All model and database files are available in the supplementary materials including an ipython notebook demonstrating all presented analyses (requires installation of ipython and COBRApy).

5 Acknowledgments

This work was supported by the Biotechnology and Biological Sciences Research Council, as part of the BBSRC Longer and Larger Grant GASCHEM [grant number BB/K00283X/1]; the BBSRC/EPSRC Synthetic Biology Research Centre Nottingham [grant number BB/L013940/1]; and the industrial partner LanzaTech Inc. Furthermore, the authors thank their colleagues at the SBRC for their support during the preparation of the paper. MP gratefully acknowledges support from Oxford Brookes University. Finally, we thank C1net for facilitating both the collaboration between Oxford Brookes University and the University of Nottingham and for funding and organising the C1net metabolic modelling workshops. The responsibility for the content of this paper lies with the authors.

6 References

[1] Armaroli, N., Balzani, V.: 'The legacy of fossil fuels', *Chem. Asian J.*, 2011, **6**, (3), pp. 768–784

[2] Pachauri, R.K., Allen, M.R., Barros, V.R., *et al.*: '*Climate change 2014: synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change*' (IPCC, Switzerland, 2014)

[3] Doran, P.M.: '*Bioprocess engineering principles*' (Academic Press, Waltham, MA, 1995)

[4] Heijstra, B.D., Leang, C., Juminaga, A.: 'Gas fermentation: cellular engineering possibilities and scale up', *Microb. Cell Fact.*, 2017, **16**, (1), p. 60

[5] Norman, R.O., Millat, T., Winzer, K., *et al.*: 'Progress towards platform chemical production using *Clostridium autoethanogenum*', *Biochem. Soc. Trans.*, 2018, **46**, (3), pp. 523–535

[6] Liew, F., Martin, M.E., Tappel, R.C., *et al.*: 'Gas fermentation – a flexible platform for commercial scale production of low carbon-fuels and chemicals from waste and renewable feedstocks', *Front. Microbiol.*, 2016, **7**, p. 694

[7] Ljungdahl, L.G., Wood, H.G.: 'Incorporation of c14 from carbon dioxide into sugar phosphates, carboxylic acids, and amino acids by *Clostridium thermoaceticum*', *J. Bacteriol.*, 1965, **89**, (4), pp. 1055–1064

[8] Ljungdahl, L.G., Wood, H.G.: 'Total synthesis of acetate from CO_2 by heterotrophic bacteria', *Annu. Rev. Microbiol.*, 1969, **23**, (1), pp. 515–538

[9] Ragsdale, S.W., Pierce, E.: 'Acetogenesis and the Wood–Ljungdahl pathway of CO_2 fixation', *Biochim. Biophys. Acta Proteins Proteomics*, 2008, **1784**, (12), pp. 1873–1898

[10] Schiel-Bengelsdorf, B., Dürre, P.: 'Pathway engineering and synthetic biology using acetogens', *FEBS Lett.*, 2012, **586**, (15), pp. 2191–2198

[11] Biegel, E., Schmidt, S., González, J.M., *et al.*: 'Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes', *Cell. Mol. Life Sci.*, 2011, **68**, (4), pp. 613–634

[12] Yoshida, M., Muneyuki, E., Hisabori, T.: 'ATP synthase – a marvellous rotary engine of the cell', *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, (9), p. 669

[13] Reidlinger, J., Müller, V.: 'Purification of ATP synthase from *Acetobacterium woodii* and identification as a na^{+}-translocating $f1f_0$-type enzyme', *Eur. J. Biochem.*, 1994, **223**, (1), pp. 275–283

[14] Köpke, M., Mihalcea, C., Liew, F., *et al.*: '2, 3-Butanediol production by acetogenic bacteria, an alternative route to chemical synthesis, using industrial waste gas', *Appl. Environ. Microbiol.*, 2011, **77**, (15), pp. 5467–5475

[15] Abrini, J., Naveau, H., Nyns, E.J.: '*Clostridium autoethanogenum*, sp. nov., an anaerobic bacterium that produces ethanol from carbon monoxide', *Arch. Microbiol.*, 1994, **161**, (4), pp. 345–351

[16] Henstra, A.M., Sipma, J., Rinzema, A., *et al.*: 'Microbiology of synthesis gas fermentation for biofuel production', *Curr. Opin. Biotechnol.*, 2007, **18**, (3), pp. 200–206

[17] Marcellin, E., Behrendorff, J.B., Nagaraju, S., *et al.*: 'Low carbon fuels and commodity chemicals from waste gases – systematic approach to understand energy metabolism in a model acetogen', *Green Chem.*, 2016, **18**, pp. 3020–3028

[18] Oberhardt, M.A., Palsson, B.Ø., Papin, J.A.: 'Applications of genome-scale metabolic reconstructions', *Mol. Syst. Biol.*, 2009, **5**, (1), p. 320

[19] Valgepea, K., de Souza Pinto Lemgruber, R., Meaghan, K., *et al.*: 'Maintenance of ATP homeostasis triggers metabolic shifts in gas-fermenting acetogens', *Cell Syst.*, 2017, **4**, (5), pp. 505–515.e5

[20] Valgepea, K., Loi, K.Q., Behrendorff, J.B., *et al.*: 'Arginine deiminase pathway provides ATP and boosts growth of the gas-fermenting acetogen *Clostridium autoethanogenum*', *Metab. Eng.*, 2017, **41**, pp. 202–211

[21] Dash, S., Ng, C.Y., Maranas, C.D.: 'Metabolic modeling of Clostridia: current developments and applications', *FEMS Microbiol. Lett.*, 2016, **363**, (4), p. fnw004

[22] Humphreys, C.M., McLean, S., Schatschneider, S., *et al.*: 'Whole genome sequence and manual annotation of *Clostridium autoethanogenum*, an industrially relevant bacterium', *BMC Genomics*, 2015, **16**, (1), p. 1

[23] Johnson, M.J., Peterson, W.H., Fred, E.B.: 'Oxidation and reduction relations between substrate and products in the acetone-butyl alcohol fermentation', *J. Biol. Chem.*, 1931, **91**, (2), pp. 569–591

[24] Brown, S.D., Nagaraju, S., Utturkar, S., *et al.*: 'Comparison of single molecule sequencing and hybrid approaches for finishing the genome of *Clostridium autoethanogenum* and analysis of CRISPR systems in industrial relevant Clostridia', *Biotechnol. Biofuels*, 2014, **7**, (1), p. 40

[25] Karp, P.D., Paley, S., Romero, P.: 'The pathway tools software', *Bioinformatics*, 2002, **18**, (suppl 1), pp. S225–S232

[26] http://sbrc-seek.nottingham.ac.uk/data_files/1?version=1

[27] Bruno-Barcena, J.M., Chinn, M.S., Grunden, A.M.: 'Genome sequence of the autotrophic acetogen *Clostridium autoethanogenum* JA1-1 strain DSM 10061, a producer of ethanol from carbon monoxide', *Genome Announcements*, 2013, **1**, (4), pp. e00628–13

[28] Köpke, M., Held, C., Hujer, S., *et al.*: '*Clostridium ljungdahlii* represents a microbial production platform based on syngas', *Proc. Natl. Acad. Sci.*, 2010, **107**, (29), pp. 13087–13092

[29] Schatschneider, S., Abdelrazig, S., Safo, L., *et al.*: 'Quantitative isotope-dilution high-resolution-mass-spectrometry analysis of multiple intracellular metabolites in *Clostridium autoethanogenum* with uniformly ^{13}C-labeled standards derived from Spirulina', *Anal. Chem.*, 2018, **90**, (7), pp. 4470–4477

[30] Caspi, R., Foerster, H., Fulcher, C.A., *et al.*: 'The Meta-Cyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases', *Nucleic Acids Res.*, 2008, **36**, (suppl 1), pp. D623–D631

[31] Fell, D.A., Poolman, M.G., Gevorgyan, A.: 'Building and analysing genome-scale metabolic models', *Biochem. Soc. Trans.*, 2010, **38**, (5), pp. 1197–1201

[32] Hartman, H.B., Fell, D.A., Rossell, S., *et al.*: 'Identification of potential drug targets in *Salmonella enterica* sv. typhimurium using metabolic modelling and experimental validation', *Microbiology*, 2014, **160**, (6), pp. 1252–1266

[33] Gevorgyan, A., Poolman, M.G., Fell, D.A.: 'Detection of stoichiometric inconsistencies in biomolecular models', *Bioinformatics*, 2008, **24**, (19), pp. 2245–2251

[34] Poolman, M.G.: 'Scrumpy: metabolic modelling with Python', *IEE Proc. Syst. Biol.*, 2006, **153**, (5), pp. 375–378

[35] Mock, J., Zheng, Y., Mueller, A.P., *et al.*: 'Energy conservation associated with ethanol formation from H_2 and CO_2 in *Clostridium autoethanogenum* involving electron bifurcation', *J. Bacteriol.*, 2015, **197**, (18), pp. 2965–2980

[36] Wang, S., Huang, H., Kahnt, J., *et al.*: 'NADP-specific electron-bifurcating [FeFe]-hydrogenase in a functional complex with for mate dehydrogenase in *Clostridium autoethanogenum* grown on CO', *J. Bacteriol.*, 2013, **195**, (19), pp. 4373–4386

[37] Schuster, S., Dandekar, T., Fell, D.A.: 'Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering', *Trends Biotechnol.*, 1999, **17**, (2), pp. 53–60

[38] Schuster, S., Hilgetag, C.: 'On elementary flux modes in biochemical reaction systems at steady state', *J. Biol. Syst.*, 1994, **2**, (2), pp. 165–182

[39] Meier, T., Ferguson, S.A., Cook, G.M., *et al.*: 'Structural investigations of the membrane-embedded rotor ring of the F_1F_O-ATPase from *Clostridium paradoxum*', *J. Bacteriol.*, 2006, **188**, (22), pp. 7759–7764

[40] Schuchmann, K., Müller, V.: 'Autotrophy at the thermo dynamic limit of life: a model for energy conservation in acetogenic bacteria', *Nat. Rev. Microbiol.*, 2014, **12**, (12), pp. 809–821

[41] Kracke, F., Virdis, B., Bernhardt, P.V., *et al.*: 'Redox dependent metabolic shift in *Clostridium autoethanogenum* by extracellular electron supply', *Biotechnol. Biofuels*, 2016, **9**, (1), p. 249

[42] Blom, J., Albaum, S.P., Doppmeier, D., *et al.*: 'EDGAR: a software framework for the comparative analysis of prokaryotic genomes', *BMC Bioinformatics*, 2009, **10**, (1), p. 154

[43] Brown, S., Santa Maria, J.P.Jr., Walker, S.: 'Wall teichoic acids of Gram-positive bacteria', *Annu. Rev. Microbiol.*, 2013, **67**, pp. 313–336

[44] Feist, A.M., Palsson, B.Ø.: 'The biomass objective function', *Curr. Opin. Microbiol.*, 2010, **13**, (3), pp. 344–349

[45] Pramanik, J., Keasling, J.: 'Stoichiometric model of *Escherichia coli* metabolism: incorporation of growth rate dependent biomass composition and mechanistic energy requirements', *Biotechnol. Bioeng.*, 1997, **56**, (4), pp. 398–421

[46] Lee, J., Yun, H., Feist, A.M., *et al.*: 'Genome-scale reconstruction and *in silico* analysis of the *Clostridium acetobutylicum* ATCC 824 metabolic network', *Appl. Microbiol. Biotechnol.*, 2008, **80**, (5), pp. 849–862

[47] Senger, R.S., Papoutsakis, E.T.: 'Genome-scale model for *Clostridium acetobutylicum*: part I. Metabolic network resolution and analysis', *Biotechnol. Bioeng.*, 2008, **101**, (5), pp. 1053–1071

[48] Senger, R.S., Papoutsakis, E.T.: 'Genome-scale model for *Clostridium acetobutylicum*: part II. Development of specific proton flux states and numerically determined sub-systems', *Biotechnol. Bioeng.*, 2008, **101**, (5), pp. 1053–1071

[49] Li, G.W., Burkhardt, D., Gross, C., *et al.*: 'Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources', *Cell*, 2014, **157**, (3), pp. 624–635

[50] Nagarajan, H., Sahin, M., Nogales, J., *et al.*: 'Characterizing acetogenic metabolism using a genome-scale metabolic reconstruction of *Clostridium ljungdahlii*', *Microbial Cell Factories*, 2013, **12**, (1), p. 1

[51] Milne, C.B., Eddy, J.A., Raju, R., *et al.*: 'Metabolic network reconstruction and genome-scale model of butanol-producing strain *Clostridium beijerinckii* NCIMB 8052', *BMC Syst. Biol.*, 2011, **5**, (1), p. 130

[52] Bainotti, A., Nishio, N.: 'Growth kinetics of *Acetobacterium* sp. on methanol-formate in continuous culture', *J. Appl. Microbiol.*, 2000, **88**, (2), pp. 191–201

[53] Oh, Y.K., Palsson, B.Ø., Park, S.M., *et al.*: 'Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high throughput phenotyping and gene essentiality data', *J. Biol. Chem.*, 2007, **282**, (39), pp. 28791–28799

[54] Hempfling, W.P., Mainzer, S.E.: 'Effects of varying the carbon source limiting growth on yield and maintenance characteristics of *Escherichia coli* in continuous culture', *J. Bacteriol.*, 1975, **123**, (3), pp. 1076–1087

[55] Verduyn, C., Postma, E., Scheffers, W.A., *et al.*: 'Energetics of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures', *Microbiology*, 1990, **136**, (3), pp. 405–412

[56] Schuster, S., Fell, D.A., Dandekar, T.: 'A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks', *Nat. Biotechnol.*, 2000, **18**, (3), p. 326

[57] Fell, D., Cornish-Bowden, A.: '*Understanding the control of metabolism*', vol. **2**, (Portland Press London, England, 1997)

[58] Cuevas, D.A., Edirisinghe, J., Henry, C.S., *et al.*: 'From DNA to FBA: how to build your own genome-scale metabolic model', *Front. Microbiol.*, 2016, **7**, p. 907

[59] Hartman, H.B.: 'Genome-scale metabolic modelling of *Salmonella* and *Lactococcus* species'. PhD thesis, Oxford Brookes University, 2013

[60] Pfeiffer, T., Sánchez-Valdenebro, I., Nuño, J., *et al.*: 'METATOOL: for studying metabolic networks', *Bioinformatics*, 1999, **15**, (3), pp. 251–257

[61] Palsson, B.Ø.: '*Systems biology: properties of reconstructed networks*' (Cambridge University Press, New York, NY, 2006)

[62] Strang, G.: '*Linear algebra and its applications*' (Thomson Learning, Boston, MA, 1988)

[63] Schellenberger, J., Que, R., Fleming, R.M., *et al.*: 'Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox v2. 0', *Nat. Protoc.*, 2011, **6**, (9), pp. 1290–1307

[64] Varma, A., Palsson, B.Ø.: 'Metabolic flux balancing: basic concepts, scientific and practical use', *Nat. Biotechnol.*, 1994, **12**, (10), p. 994
[65] Schuetz, R., Kuepfer, L., Sauer, U.: 'Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*', *Mol. Syst. Biol.*, 2007, **3**, (1), p. 119
[66] Holzhütter, H.G.: 'The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks', *Eur. J. Biochem.*, 2004, **271**, (14), pp. 2905–2922
[67] Holzhütter, H.G.: 'The generalized flux-minimization method and its application to metabolic networks affected by enzyme deficiencies', *Biosystems*, 2006, **83**, (2–3), pp. 98–107
[68] Ebrahim, A., Lerman, J., Palsson, B.*, et al.*: 'COBRApy: constraints-based reconstruction and analysis for Python', *BMC Syst. Biol.*, 2013, **7**, (1), p. 74
[69] Machado, D., Andrejev, S., Tramontano, M.*, et al.*: 'Fast automated reconstruction of genome-scale metabolic models for microbial species and communities', *Nucleic Acids Res.*, 2018, **46**, (15), pp. 7542–7553
[70] Cotter, J.L., Chinn, M.S., Grunden, A.M.: 'Ethanol and acetate production by *Clostridium ljungdahlii* and *Clostridium autoethanogenum* using resting cells', *Bioprocess. Biosyst. Eng.*, 2009, **32**, (3), pp. 369–380
[71] Breitkopf, R.: 'Understanding the C_4 dicarboxylic acid metabolism in *Clostridium autoethanogenum*'. PhD thesis, University of Nottingham, 2018
[72] Valgepea, K., Lemgruber, R.S.P., Abdalla, T.*, et al.*: 'H_2 drives metabolic rearrangements in gas-fermenting *Clostridium autoethanogenum*', *Biotechnol. Biofuels*, 2018, **11**, (1), p. 55
[73] Angenent, L.T., Karim, K., Al-Dahhan, M.H.*, et al.*: 'Production of bioenergy and biochemicals from industrial and agricultural wastewater', *Trends Biotechnol.*, 2004, **22**, (9), pp. 477–485
[74] Claassen, P., Van Lier, J., Contreras, A.L.*, et al.*: 'Utilisation of biomass for the supply of energy carriers', *Appl. Microbiol. Biotechnol.*, 1999, **52**, (6), pp. 741–755
[75] Levin, D.B., Pitt, L., Love, M.: 'Biohydrogen production: prospects and limitations to practical application', *Int. J. Hydrog. Energy*, 2004, **29**, (2), pp. 173–185
[76] Le Novère, N., Bornstein, B., Broicher, A.*, et al.*: 'Biomodels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems', *Nucleic Acids Res.*, 2006, **34**, (suppl_1), pp. D689–D691